A.A. Orlov [1,2]
A.Ya. Zherebker [2]
A.A. Eletskaya [1]
V.S. Chernikov [1]
L.I. Kozlovskaya [1]
I.V. Perminova [2]
D.I. Osolodkin [1,2]

EXPLORING CHEMICAL SPACE OF HUMIC SUBSTANCES THROUGH ESI FTICR MS ANALYSIS AND SIMILARITY SEARCHING

[1] Chumakov Institute of Poliomyelitis and Viral Encephalitides, Chumakov FSC R&D IBP RAS, prem. 8 korp. 1, pos. Instituta Poliomielita, pos. Moskovsky, Moscow 108819, Russia.
[2] Department of Chemistry, Lomonosov Moscow State University, 1 bd. 3 Leninskie Gory, Moscow 119991, Russia

rnase@qsar.chem.msu.ru

Microbial and abiotic degradation of biomacromolecules leads to formation of heterogenous supramolecular mixtures of natural organic compounds, named humic substances (HS). HS are the major organic constituents of soils, coals, natural waters, etc. They have a broad spectrum of biological activity, such as anti-inflammatory, antibacterial, and antiviral. Activity of HS against a wide range of enveloped and non-enveloped DNA and RNA viruses allows to consider them as a promising tool against viral pathogens with no approved therapeutics. The structure and composition of HS as well as the mechanism of their action are unclear. Analysis of HS components by high-resolution mass-spectrometry revealed a large diversity of the molecular formulas of HS [1]. These formulas are attributed to the most common classes of natural organic compounds such as lignins, tanins, terpenoids, flavonoids, etc., which could act through interfering with different stages of virus life cycle [2]. In this study we proposed a methodology for identification of active HS components using Fourier transform ion cyclotron resonance mass-spectrometry (FTICR MS) and chemical database search. This methodology was applied for a study of HS antiviral activity against tick-borne encephalitis virus (TBEV) and enteroviruses.

Nine HS samples obtained by standard protocol were tested against TBEV and enteroviruses reproduction in cell-based assays. All of them inhibited TBEV reproduction with EC_{50} values of approximately 1 μg/ml and exhibited no cytotoxicity up to 10 μg/ml. The same samples did not inhibit reproduction of enteroviruses in the similar concentrations. The molecular composition of the samples was determined by ESI FTICR MS in negative ion mode of electrospray. The analysis of the spectral data and calculated molecular formulas using principal component analysis (PCA), self-organising maps (SOM) and Van-Krevelen plots revealed the formulas, to which the activity of the samples could be attributed. Similarity search for these formulas in large bioactivity databases (ChEMBL, PubChemBioAssay) was further used to find structures with relevant activity profile. Several dozens unique structures were retrieved for each sample. These data were standardised and properly annotated to ICTV virus taxonomy. Thus, a database of formulas, structures, bioactivities was created. Chemical space was visualised using PCA, SOM, and RBS, and specific features of the compounds were characterised. Scaffold analysis revealed polyphenols and saponin analogs to be the most represented. Several compounds from our database were previously tested on activity against dengue virus NS3 protease. Taking into account the similarity between dengue and TBEV proteases, we suggested that inhibition of TBEV NS3 protease could be an explanation of anti-TBEV activity of HS. Despite the exact mechanism of action remains to be studied, our approach allows making presumptions about the mode of action and the structures of the compounds behind HS antiviral activity.

1. Zherebker A.Ya. et.al. *Mendellev Communication*, 2016, **26**: 446-448.
2. Zhernov Y.V. et al. *New Journal of Chemistry*, 2017, **41**: 212-224.

This work was supported by the Russian Foundation for Basic Research (project 16-03-01057).